

APPARATUS FOR COMPRESSING AND EXPANDING
A DIGITAL INPUT SIGNALBACKGROUND OF THE INVENTIONField of the Invention

This invention relates to an apparatus that applies block floating processing to a digital input signal, orthogonally transforms the block floating processed signal on the time axis into plural spectral coefficients on the frequency axis, divides the spectral coefficients into plural critical bands, and carries out adaptive bit allocation to quantize the spectral coefficients in each critical band.

Description of the Prior Art

As one of technologies for compressing a digital audio signal, and similar analog signals, it is known to apply block floating processing in which the digital input signal is divided into blocks of a predetermined number of words, and block floating processing is applied to each block. In known block floating processing, the maximum one of the absolute values of the words in the block is sought, and is used as a common block floating coefficient for all the words in the block.

Further, it is also known to use orthogonal transform coding to transform orthogonally a signal on the time axis into a signal on the frequency axis. The resulting spectral coefficients are then quantized. For example, it is known to divide, e.g., a PCM audio signal into blocks, each of a predetermined number of words, and to apply a Discrete Cosine Transform (DCT) to each block. In addition, it is also known to divide the spectral coefficients resulting from an orthogonal transform into critical bands and to quantize the spectral coefficients by applying adaptive bit

#3M 08/159125
Sub spec B
3794B1
C2195
10

allocation to each critical band. The number of bits allocated for quantizing the spectral coefficients in each critical band is determined depending on an allowable noise level in each critical band, which takes psychoacoustic masking into consideration.

5 There are many instances where the operational processing for an orthogonal transform is executed by using a FIR (Finite-duration Impulse-Response) filter of the multi-tap type. This type of operational processing includes coefficient multiplication processing and/or operations for calculating a sum total, etc. The number of bits generated by such
10 processing results the likelihood of overflows. To prevent such overflows, the number of bits generated by the operation must be allowed for in advance by, e.g., processing using several orders of bits greater than the number of bits in each word of the input signal. For such a multi-bit operation, a high performance DSP (Digital Signal Processing unit) is required, and it takes much time as well. Accordingly, simplification of the orthogonal transform processing is desirable.

BI
COOT
15
20 In view of this, a technique has been proposed to apply the above-mentioned block floating processing to the digital input signal prior to the orthogonal transform processing. The block floating processing achieves bit compression of the input signal and reduces the number of bits subject to the orthogonal transform operation.

25 Further, a technique has been also proposed to adaptively vary the size of the block subject to the orthogonal transform processing depending on a signal. Such a technique is employed because, particularly when the input signal is divided into components in several (e.g., about three) frequency ranges, and the orthogonal transform processing is performed in each frequency range, varying the block length in response to the magnitude of temporal changes, or in response to a pattern, etc., in the frequency range
30 signals permits a more efficient quantizing of the resulting spectral components than when the block length is fixed.

3

It is to be noted that when block floating is applied prior to the orthogonal transform processing, and the block length is adaptively changed depending on a signal, independent processing is applied, which results in the drawback that the amount of processing required is increased.

5 For example, as shown in Fig. 15, a relatively large block BL is divided in advance into several sub blocks (e.g., the four sub blocks BL_{S1} , BL_{S2} , BL_{S3} and BL_{S4}). As indicated by step S31 of Fig. 16, the respective energies of the sub blocks BL_{S1} , BL_{S2} , BL_{S3} and BL_{S4} are calculated in the process of determining the size of the variable length block. At the next step
10 S32, the block size is determined in response to the energies of the respective sub blocks. Then, at step S33, the maximum absolute value within the block determined in the previous step is calculated to implement block floating processing using the calculated maximum absolute value. At the next step S34, orthogonal transform processing, such as DCT, is applied to the block.

In such a processing procedure, calculation of the energy of each respective sub block BL_{S1} , BL_{S2} , BL_{S3} and BL_{S4} for determining the block size, and calculation of the maximum absolute values in the thus-determined blocks for applying the block floating processing are required. As a result,
20 the quantity subject to processing or the number of steps in processing by a so-called microprogram is increased.

When determining an allowable noise level for each critical band to take account of masking, it has been proposed to correct the allowable noise level to take into consideration the minimum audible level characteristic of
25 the human sense of hearing. In this, an allowable noise level already calculated is compared with a minimum audible level, and the greater level is selected as the new allowable noise level.

The allowable noise level in which masking is taken into consideration is assumed to be constant across each critical band. However,
30 since the minimum audible level is measured using a sine wave, the can be

BL
cont
15

an appreciable change in the minimum audible level between the low frequency end and the high frequency end of each critical band. This is particularly so at high frequencies, where the critical bands are relatively broad. For this reason, using a single minimum audible level for each critical band causes appreciable errors, resulting in the possibility of an excess number of bits being allocated for quantizing the spectral coefficients towards the high frequency end of the critical band.

In addition, although it is conceivable to divide the critical band into small sub bands, and to give a minimum audible level for each sub band, this is not preferable because the quantity of information required to be transmitted is increased.

SUMMARY OF THE INVENTION

This invention has been proposed, and its object is to provide an apparatus for compressing a digital input signal. Block floating is applied prior to the orthogonal transform processing and the length of the block subject to transform processing is changed depending on a signal. The apparatus is constructed so that the quantity subject to processing is reduced.

Another object of this invention is to provide an apparatus for compressing a digital input signal in which, when the input signal is divided in frequency into spectral coefficients in critical bands, and adaptive bit allocation is applied thereto on the basis of allowable noise levels, errors in the minimum audible level are reduced in those critical bands in which the minimum audible level is selected as the allowable noise level.

Accordingly, a first aspect of the invention provides an apparatus for compressing a digital input signal. The apparatus comprises an index generating circuit that generates an index in response to the digital input signal. Also included in the apparatus are a block length decision circuit, which determines a division of the digital input signal into blocks in response to the index, and a block floating processing circuit, which applies block

B1
cont

floating processing to the blocks of the digital input signal in response to the index. The circuit further includes an orthogonal transform circuit that orthogonally transforms each block floating processed block of the digital input signal to produce plural spectral coefficients. Finally, the circuit

5 comprises an adaptive bit allocation circuit that divides the plural spectral coefficients into bands, and adaptively allocates a number of quantizing bits to quantize the spectral coefficients in each of the bands.

A variation of the first aspect of the invention provides an apparatus for compressing a digital input signal. The apparatus comprises a band

10 division filter that divides the digital input signal into a frequency range signal in each of plural frequency ranges. Also included in the apparatus are a block length decision circuit, which determines a division of each frequency range signal in time into blocks in response to an index, and a block floating processing circuit, which applies block floating processing to each frequency range signal in response to the index. The circuit also

15 includes an orthogonal transform circuit that orthogonally transforms each block floating processed frequency range signal to produce plural spectral coefficients. The orthogonal transform circuit transforms each frequency range signal in blocks determined by the block length decision means.

20 Finally, the apparatus comprises an adaptive bit allocation circuit that divides the plural spectral coefficients into bands, and adaptively allocates numbers of quantizing bits for quantizing the spectral coefficients in response to an allowable noise level in each of the bands.

A second aspect of the invention provides an apparatus for

25 compressing a digital input signal. The apparatus comprises a circuit that derives plural spectral coefficients from the digital input signal, and an adaptive bit allocation circuit that divides the spectral coefficients by frequency into bands, and adaptively allocates a number of quantizing bits for quantizing the spectral coefficients in each band in response to an

30 allowed noise level for each of the bands. The adaptive bit allocation circuit

BI
cont

includes an allowable noise level calculation circuit that calculates an allowed noise level for each band, a comparator that compares the allowable noise level with a minimum audible level in each band, and a selector that selects the minimum audible level as the allowable noise level for each band in which the comparator determines that the minimum audible level is higher than the allowable noise level.

5 A variation on the second aspect of the invention provides an apparatus for compressing a digital input signal. The apparatus comprises a band division filter that divides the digital input signal into a frequency range signal in each of plural frequency ranges. Also included in the apparatus are
10 a block floating processing circuit that applies block floating processing to each frequency range signal divided in time into blocks, and an orthogonal transform circuit that orthogonally transforming each block of each frequency range signal to provide plural spectral coefficients. Finally, the apparatus includes an adaptive bit allocation circuit that divides the spectral coefficients into bands, and adaptively allocates a number of quantizing bits for quantizing the spectral coefficients in each band in response to an allowable noise level in each band. The adaptive bit allocation circuit
15 includes an allowable noise level calculation circuit that calculates the allowable noise level for each band, and a comparator for comparing the allowable noise level with a minimum audible level in each band, and that sets a flag for each band in which the minimum audible level is higher than the allowable noise level. Finally, the adaptive bit allocation circuit includes
20 a selector that selects the minimum audible level as the allowed noise level in each band in which the flag is set.

25 When the allowable noise level in each critical band is determined by the minimum audible level, bit allocation is carried out according to the allowable noise level in plural sub bands obtained by further dividing the critical band in frequency. When this is done, a flag indicating that the
30 minimum audible level has been adopted as the allowable noise level for the

BI
cont

band only needs to be transmitted. This avoids the necessity of transmitting allowable noise level information for each sub band. Accordingly, accurate allowable noise levels can be provided without increasing the quantity of auxiliary information transmitted. This provides an improvement in signal quantity without degrading the signal compression efficiency. In addition, even if the absolute value of the minimum audible level is altered later, compatibility can be maintained.

A third aspect of the invention provides a method for compressing a digital input signal. In the method, an index is generated in response to the digital input signal, a division of the digital input signal into blocks is determined in response to the index, and block floating processing is applied to the blocks of the digital input signal in response to the index. Each block floating processed block of the digital input signal is orthogonally transformed to produce plural spectral coefficients, the spectral coefficients are divided into bands, and numbers of quantizing bits are adaptively allocated to quantize the spectral coefficients in each band.

A fourth aspect of the invention provides a method for compressing a digital input signal. In the method, plural spectral coefficients are derived from the digital input signal, the spectral coefficients are divided by frequency into bands, and a number of quantizing bits is allocated for quantizing the spectral coefficients in each band in response to an allowed noise level for each band. In the step of adaptively allocating a number of quantizing bits, an allowable noise level is calculated for each band, the allowable noise level is compared with a minimum audible level in each band, and the minimum audible level is selected as the allowable noise level in each band in which the minimum audible level is higher than the allowable noise level.

A fifth aspect of the invention provides an apparatus for expanding a compressed digital signal. The compressed digital signal includes plural quantized spectral coefficients and auxiliary information. The apparatus

BI
CONT
15

comprises adaptive bit allocation decoding circuit that operates in response to the auxiliary information and inversely quantizes the quantized spectral coefficients to provide plural spectral coefficients. The circuit also includes a block floating circuit that applies block floating to the spectral coefficients. Also included in the apparatus is an inverse orthogonal transform circuit means that inversely orthogonally transforms the block floating processed spectral coefficients to provide plural frequency range signals. Finally, the apparatus includes an inverse filter circuit that synthesizes the frequency range signals to provide an output signal.

B1
Co¹⁵
Cat

A sixth aspect of the invention provides a method for expanding a compressed digital signal to provide a digital output signal. The compressed digital signal includes plural quantized spectral coefficients divided by frequency into bands. At least one of the bands is a divided band in which the spectral coefficients in the band are further divided by frequency into sub bands. The compressed digital signal additionally includes an allowed noise level for each band, and, for each divided band, a flag signal, The quantized spectral coefficients in each band and sub band are quantized using an adaptively-allocated number of quantizing bits. In the method, in each divided band, the allowed noise level of the band is set as the allowed noise level for the band when the flag signal for the band is in a first state. Also, in each divided band, the allowed noise level of the band is set as the allowed noise level for one of the sub bands constituting the band when the flag signal for the band is in a second state. Finally, in each divided band, an allowed noise level for each of the other sub bands constituting the band is calculated from the allowed noise level of the band. The allowable noise level for each band and sub band is then used to inversely quantize the respective quantized spectral coefficients in each band and sub band, and the digital output signal is derived from the resulting spectral coefficients.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a circuit diagram showing, in a block form, the outline of the configuration of an apparatus for compressing a digital input signal according to an embodiment of this invention.

5 Fig. 2 is a view showing an actual example of how the input signal is divided into frequency ranges and how the input signal is divided in time into blocks in each frequency range in the embodiment.

Fig. 3 is a flow chart for explaining the essential part of the process by which the allowable noise level is set in the embodiment.

10 Fig. 4 is a view showing a critical band used for explaining how the allowable noise level is set in the embodiment.

Fig. 5 is a flow chart for explaining the essential part of the decoding operation in the embodiment.

Fig. 6 is a view showing a critical band used for explaining the decoding operation in the embodiment.

Fig. 7 is a view showing an example in which the block size in one frequency range is switched between two sizes in the apparatus of Fig. 1.

Fig. 8 is a view showing an example in which the block size in one frequency range is switched between three sizes in the apparatus of Fig. 1.

20 Fig. 9 is a flow chart for explaining the block floating operation of the embodiment.

Fig. 10 is a circuit diagram showing, in a block form, an actual example of allowable noise calculation circuit 20 of the apparatus shown in Fig. 1.

25 Fig. 11 is a view showing a bark spectrum.

Fig. 12 is a view showing a masking spectrum.

Fig. 13 is a view in which a minimum audible level curve and a masking spectrum are synthesized.

Fig. 14 is a block diagram showing an actual example of a decoder to which the embodiment of this invention can be applied.

B1
cont
15

Fig. 15 is a view showing an example of the length of a block by the processing procedure in the prior art.

Fig. 16 is a flow chart showing an example of the procedure of a conventional block floating processing.

5

DESCRIPTION OF THE PREFERRED EMBODIMENT

A preferred embodiment of this invention will now be described with reference to the attached drawings.

This invention can be applied to an apparatus for compressing a digital input signal, such as a PCM audio signal, etc., using subband coding (SBC), adaptive transform coding (ATC), and adaptive bit allocation (APC-AC). In the apparatus of the embodiment shown in Fig. 1, the digital input signal is divided into frequency range signals in plural frequency ranges. The bandwidths of the frequency ranges increase with increasing frequency. Orthogonal transform processing is applied to each frequency range signal to provide plural spectral coefficients. Adaptive bit allocation is used to quantize the spectral coefficients divided into critical bands in which the masking characteristic of the human sense of hearing is taken into consideration. In addition, in the embodiment of this invention, the block size is adaptively varied in response to a signal prior to the orthogonal transform processing, and block floating processing is applied to every block.

In Fig. 1, input terminal 10 is supplied with a PCM audio signal in the frequency range of, e.g., 0 Hz to 20 kHz. This input signal is divided into a signal in the frequency range of 0 Hz to 10 kHz and a frequency range signal in the frequency range of 10 to 20 kHz by using a band division filter 11, e.g., a Quadrature Mirror Filter (QMF filter), etc. The signal in the frequency range of 0 Hz to 10 kHz is further divided into a frequency range signal in the frequency range of 0 Hz to 5 kHz and a frequency range signal in the frequency range of 5 to 10 kHz by using a band division filter

B1
cont

12, e.g., a QMF filter, etc. The signal in the frequency range of 10 to 20 kHz from the band division filter 11 is sent to the Discrete Cosine Transform (DCT) circuit 13 serving as an orthogonal transform circuit, the signal in the frequency range of 5 to 10 kHz from the band division filter 12 is sent to the DCT circuit 14, and the signal in the frequency range of 0 Hz to 5 kHz from the band division filter 12 is sent to the DCT circuit 15. Thus, these signals are subjected to DCT processing, respectively.

In the embodiment of this invention, in order to reduce the quantity of operations in the orthogonal transform processing, block floating processing is applied to the frequency range signals prior to the orthogonal transform processing. This provides data compression. The block floating is released after the block floating processed signals have been orthogonally transformed.

In Fig. 1, the frequency range signals obtained from the band division filters 11 and 12 are delivered to a block floating processing circuit 16, in which block floating processing is carried out using the respective blocks BL as shown in Fig. 15. In the transform circuits (DCT, i.e., Discrete Cosine Transform, circuits are shown in the example of Fig. 1) 13, 14 and 15, orthogonal transform processing is applied to the signals which have undergone such block floating processing. Thereafter, the block floating is released by the block floating release circuit 17. In releasing the block floating, block floating information from the block floating processing circuit 16 is used. Block floating coefficients may be determined in the block floating processing by taking the logical sum of the absolute values of the words in each block.

Fig. 2 shows an actual example of how a frame of the digital input signal is divided into blocks in each frequency range prior to delivery to the DCT circuits 13, 14 and 15. In the actual example of Fig. 2, the bandwidth of the frequency ranges increases and the time resolution increases (i.e., the block size is reduced) as the frequency increases. For the frequency range

BL
out

signal in the low frequency range of 0 Hz to 5 kHz, one block B_{LL} is chosen to have, e.g., 1024 samples. For the frequency range signal in the middle frequency range of 5 to 10 kHz, the frame is divided into two blocks B_{LM1} and B_{LM2} , each having a length $T_{BL}/2$, one half of the length T_{BL} of the block B_{LL} in the low frequency range. For the frequency range signal in the high frequency range of 10 to 20 kHz, the signal is divided into four blocks B_{LH1} , B_{LH2} , B_{LH3} and B_{LH4} , each having a length $T_{BL}/4$, one fourth of the length T_{BL} of the block B_{LL} of the low frequency range. It is to be noted that, in the case where the input signal has a frequency range of 0 Hz to 22 kHz, the low frequency range extends from 0 Hz to 5.5 kHz, the middle frequency range extends from 5.5 to 11 kHz, and the high frequency range extends from 11 to 22 kHz.

BI
cont
15
In the embodiment of this invention, as will be described later, the block size is caused to vary in response to a signal, and determination of the block size is carried out in response to the maximum absolute value used also for determining the block floating coefficients of the block floating.

Turning back to Fig. 1, the spectral coefficients obtained as the result of the DCT processing in the respective DCT circuits 13, 14 and 15, are subject to block floating release processing in the block floating release circuit 17, and are then divided by frequency into critical bands. The spectral coefficients are sent to the adaptive bit allocation circuit 18.

20
A critical band is division of the frequency range that takes into account characteristics of the human sense of hearing. A critical band is the band of noise that can be masked by a pure signal that has the same intensity as the noise and has a frequency in the middle of the critical band. The bandwidth of successive critical bands increases with increasing frequency. The audio frequency range of 0 Hz to 20 kHz is normally divided into, e.g., 25 critical bands.

25
30
The allowable noise calculation circuit 20 calculates an allowable noise level for each critical band, taking into account the masking effect.

The spectral coefficients are divided into plural critical bands to calculate the number of bits to be allocated to quantize the spectral coefficients in each critical band. Quantizing bits are allocated on the basis of the allowable noise level and the energy or peak value, etc. in each critical band. In response to the bit numbers allocated to each critical band by the adaptive bit allocation circuit 18, the spectral coefficients are quantized. The quantized spectral coefficients are taken out through the output terminal 19.

The allowable noise calculation circuit 20 is supplied with a minimum audible level for each critical band from a minimum audible level curve generator 32. Each minimum audible level is compared with the allowable noise level, in which the masking effect is taken into consideration, in the comparator 35. As a result, when the minimum audible level is higher than the allowable noise level, the minimum audible level is selected as the allowable noise level.

According to the invention, some critical bands, especially the higher frequency critical bands, are divided into sub bands to take into consideration the error in the minimum audible level that occurs particularly in critical bands having a wide bandwidth. Dividing critical bands allows a minimum audible level for each sub band to be used for the respective allowable noise level for each sub band. Bit allocation is then carried out for each sub band.

The operation of the division into sub bands will now be described with reference to Figs. 3 and 4.

Fig. 3 is a flow chart for explaining the operation, and Fig. 4 shows the example where one critical band B is divided into sub bands BB (four sub bands are shown in the example of Fig. 4).

In the step S1 of Fig. 3, it is determined whether or not the level of the minimum audible level curve RC of the sub band BB_1 , the lowest frequency sub band of the four sub bands BB_1 to BB_4 of the critical band B, is higher than the masking level, which is the present allowable noise

BI
cont

determined in consideration of masking ($RC > MS$). If the result of this step S1 is YES (the level of the minimum audible level curve RC is higher than the masking level MS), the operation proceeds to step S2 where the minimum audible level is selected as the allowable noise level. The flag F_{RC} is set at the next step S3 ($F_{RC}=1$). The operation proceeds to step S4 where adaptive bit allocation is carried out using the level of the minimum audible level curve RC as the allowable noise level. Conversely, when the result of the step S1 is NO, the operation proceeds to step S5 where the masking level is selected as the allowable noise level. The flag F_{RC} is cleared to 0 at step S6, and the process reverts to the step S4 where adaptive bit allocation is carried out.

Various possibilities for one critical band B are shown in Fig. 4. Where the minimum audible level curve is the curve RCa and the masking level is the level MS, the result of the step S1 is YES. Where the minimum audible level curve is the curve RCb or the curve RCc, the result of the step S1 is NO. When the minimum audible level curve is the curve RCa, the minimum audible level curve RCa is selected as the allowable noise level, and bit allocation is carried out in each sub band BB_1 to BB_4 in response to the allowable noise level in each sub band BB_1 to BB_4 . On the other hand, when the minimum audible level curve is the curve RCb or the curve RCc, the masking level MS is selected as the allowable noise level, and bit allocation is carried out in response to a single allowable noise level throughout the whole critical band B.

The allowable noise level for each critical band is transmitted from the compressor as auxiliary information, along with quantized spectral coefficients as main information. This is so, even when the minimum audible level curve RCa is selected as the allowable noise level. The auxiliary information transmitted is a single allowable noise level for each critical band. The minimum audible level curve is determined from the characteristics of the human sense of hearing. Thus, a minimum audible

B1
cont
15

level curve pattern, or relative value data, etc., can be stored in advance into a ROM, etc. The minimum audible level of the other sub bands BB_2 to BB_4 can easily be determined from the data in the ROM in response to the minimum audible level of, e.g., the lowest-frequency sub band BB_1 .

5 Fig. 5 is a flow chart for explaining an essential part of the expansion processing in a complementary expander. At step S11 of Fig. 5, it is determined whether or not the flag F_{RC} is 1. If the result is YES, i.e., the allowable noise level of the corresponding critical band is given by the minimum audible level curve, the allowable noise level for each sub band BB_1 to BB_4 is calculated at the next step S12. Even though only one
10 allowable noise level is transmitted for all of the critical band B, e.g., the allowable noise level NL_1 of the lowest frequency sub band BB_1 , as shown in Fig. 6, allowable noise levels NL_2 to NL_4 for the other sub bands BB_2 to BB_4 can be determined by calculation from the pattern of the minimum audible level curve RC by making use of a relative list, etc. of minimum audible
15 level values stored in a ROM, etc. as described above.

If the result at the step S11 is NO, i.e., the allowable noise level for the critical band is given by the masking level MS, the operation proceeds to step S13 where a fixed allowable noise level is set for the whole of the
20 critical band B. Bit allocation decoding processing takes place at step S14 in response to the allowable noise level determined at the respective one of the steps S12 and S13.

The method of determining the division of each frame of each frequency range signal into the blocks in which the respective frequency range signals are orthogonally transformed, and the way in which the block
25 sizes are adaptively changed in response to a signal, i.e., each respective frequency range signal, will now be described.

First, the case where the division of the frame is switched between a block BL, having a block length T_{BL} , and two blocks BL_{R1} and BL_{R2} each
30 having a block length of $T_{BL}/2$, one half of T_{BL} , will be described with

reference to Fig. 7. First, maximum absolute values (or logical sums) MX_{R1} and MX_{R2} in the respective sub blocks corresponding to the smaller blocks BL_{R1} and BL_{R2} are determined. Then, these maximum absolute values MX_{R1} and MX_{R2} are compared. When the ratio therebetween is as indicated by the following equation (1), the frame is divided into the smaller blocks BL_{R1} and BL_{R2} .

$$MX_{R2}/MX_{R1} \geq 20 \quad \dots (1)$$

Otherwise, a block size equal to the size of the larger block BL is selected.

Next, the case where the division of the frame is switched between a large block BL having a block length T_{BL} , medium blocks BL_{R1} and BL_{R2} having a block length $T_{BL}/2$, one half of the block length T_{BL} , and small blocks BL_{S1} , BL_{S2} , BL_{S3} and BL_{S4} , each having a block length $T_{BL}/4$, one fourth of the block length T_{BL} , will be described with reference to Fig. 8. First, respective maximum absolute values (or logical sums) MX_{S1} , MX_{S2} , MX_{S3} and MX_{S4} in the sub blocks corresponding to the small blocks BL_{S1} , BL_{S2} , BL_{S3} and BL_{S4} are determined. With respect to these four maximum absolute values MX_{S1} , MX_{S2} , MX_{S3} and MX_{S4} , if the following relationship indicated by the following equation (2) holds, the frame is divided into blocks equal to the small blocks BL_{S1} , BL_{S2} , BL_{S3} and BL_{S4} , having a length of $T_{BL}/4$.

$$MX_{Sn+1}/MX_{Sn} \geq 20 \quad \dots (2)$$

where n is 1, 2 or 3.

If the above equation (2) is not satisfied, the respective maximum absolute values (or logical sums) MX_{R1} and MX_{R2} in the sub blocks corresponding to the medium blocks BL_{R1} and BL_{R2} are determined. Then it is determined whether or not the following equation (3) is satisfied.

$$MX_{R2}/MX_{R1} \geq 10 \quad \dots (3)$$

If the above equation (3) is satisfied, the frame is divided into blocks equal to the medium blocks BL_{R1} and BL_{R2} , having a length $T_{BL}/2$. Otherwise,

BL
cont

i.e., the following equation (4) holds, and the frame remains undivided, with a block size of the large block BL having a length T_{BL} .

$$MX_{R2}/MX_{R1} < 10 \quad \dots (4)$$

Fig. 9 shows a software routine for processing each frequency range signal prior to orthogonal transform processing. Each frequency range signal comprises plural words. In Fig. 9, at step S111, the absolute value of each word is first calculated. At the next step S112, the maximum absolute value is detected. Instead of detecting the maximum absolute value, a logical sum operation may be performed. At the next step S113, it is determined whether the maximum absolute value of all the words in the sub block or whether the logical sum of all the words in the sub block has been taken. The sub block is an integral fraction of the frame, for instance one half (first example above) or one fourth (second example above) of the frame. When it is determined at the step S113 that the logical sum operation (or the absolute maximum value determination) of all the words is not completed (NO), the operation returns to the step S111. On the other hand, when the logical sum operation (or absolute maximum determination) of all the words is completed (YES), the operation proceeds to the next step S114.

At step S114, if the logical sum of the absolute values in the sub block is taken at step S112, processing to detect the maximum absolute value in the sub block is unnecessary. Floating coefficients (shift quantities) can be determined by simple processing including only a logical sum operation.

The steps S114 and S115 provide the operation for determining the shift quantity as the block floating coefficient. At the step S114, a left shift is carried out. At the step S115, it is determined whether the Most Significant Bit (MSB) of the shift result is equal to "1." If a "1" is not detected as the MSB at the step S115 (NO), the operation returns to the step S114. Otherwise, if a "1" is detected (YES), the operation proceeds to the next step S116.

BL
out

At the step S116, it is determined whether the maximum absolute value (or shift quantity) of all sub blocks of the different sizes has been obtained. When the result is NO, the operation returns to the step S111. Otherwise, if the result is YES, the operation proceeds to the next step S117.

5 At the step S117, the block size is determined using the above equation (1) or the above equations (2) to (4), and the maximum absolute value (or logical sum) of the block thus determined is calculated. At the next step S119, the words in the determined block are normalized (i.e., are subjected to floating processing). At step S120, it is determined whether all the words
10 in the determined block have been normalized. If the result is NO, the operation returns to the step S119. Otherwise, if the result is YES, the operation proceeds to the next step S121. At the step S121, it is determined whether, when, e.g., the block size of the medium blocks BL_{R1} and BL_{R2} or the small blocks BL_{S1} , BL_{S2} , BL_{S3} and BL_{S4} , is selected, the processing with respect to all blocks in the frame has been completed. If the result is NO,
15 the operation returns to the step S111. Otherwise, if the result is YES, the operation proceeds to the next step S122. At the step S122, the orthogonal transform processing is carried out. The processing is thus completed.

20 In accordance with this embodiment, by using the maximum absolute value (or logical sum) calculated for each block to determine both the block floating coefficient and the block size, the quantity subject to processing can be reduced. Thus, the number of steps, when, e.g., processing is carried out using a microprogram, can be reduced.

25 Fig. 10 is a circuit diagram showing, in block form, the outline of the configuration of an actual example of the allowable noise calculation circuit 20. In Fig. 10, input terminal 21 is supplied with spectral coefficients from the respective DCT circuits 13, 14 and 15. An amplitude value and a phase value are calculated from the real number component and the imaginary number component of each spectral coefficient. This approach
30 is employed in consideration of the fact that the human sense of hearing is

Bl
cont

considerably more sensitive in the frequency domain to amplitude than to phase.

The resulting amplitude values in the frequency domain are sent to the energy calculation circuit 22 in which an energy for each critical band is determined by, e.g., calculating the sum total of the respective amplitude values in the critical band, or any other appropriate method. Instead of determining the energy in each critical band, there are instances where a peak value, or a mean value of the amplitude values in the band may be used. The output from the energy calculation circuit 22, e.g., a spectrum of the energy sum in each respective critical band is generally called a bark spectrum. Fig. 11 shows such a bark spectrum SB for each critical band. To simplify the figure, only twelve bands (B_1 to B_{12}) are shown.

To allow for the influence of the masking of the bark spectrum SB, convolution processing is implemented to multiply the bark spectrum SB by predetermined filter coefficients and to add the multiplied results. To realize this, the output from the energy calculation circuit 22 in each critical band, i.e., respective values of the bark spectrum SB, is sent to the convolution filter circuit 23. This convolution filter circuit 23 comprises, e.g., plural delay elements for sequentially delaying input data, plural multipliers (e.g., 25 multipliers, one for each critical band) for multiplying the outputs from the delay elements by filter coefficients, and a sum total adder for summing the multiplier outputs. By this convolution processing, the sum total of the portion indicated by dotted lines in Fig. 11 is calculated.

Masking is a psychoacoustic phenomenon in which a signal is rendered inaudible if it is masked by another signal. There is temporal masking, in which a signal is masked by a signal occurring before or after it in time. There is also simultaneous masking, in which a signal is masked by a simultaneously-occurring signal of a different frequency. As a result of masking, if there is any noise in a portion of the spectrum subject to masking, such noise will be inaudible. For this reason, with an actual audio

B1
cont
15

signal, any noise within the masking range of the signal is inaudible, and is regarded as allowable noise.

An actual example of the filter coefficients of the respective multipliers of the convolution filter circuit 23 will now be described.

5 Assuming that the coefficient of a multiplier M corresponding to an arbitrary band is 1, the multiplying operation is carried out as follows: at the multipliers M-1, M-2, M-3, M+1, M+2, and M+3, the outputs from the respective delay elements are multiplied by the filter coefficients of 0.15, 0.0019, 0.0000086, 0.4, 0.06, and 0.007, respectively. Thus, convolution
10 processing of the bark spectrum SB is carried out. M is an arbitrary integer of 1 to 25.

31
cont
15 The output of the convolution filter circuit 23 is sent to a subtractor 24. The subtractor 24 determines the level α corresponding to the allowable noise level in the convoluted region. The level α is the level that gives an allowable noise level for each critical band by deconvolution as will be described below. An allowed function (i.e., a function representing the masking level) for determining the level α is delivered to the subtractor 24. By increasing or decreasing this allowed function, control of the level α is carried out. The allowed function is delivered from the $(n - ai)$ function
20 generator 25, which will be described later.

The level of α corresponding to the allowable noise level is determined by the following equation:

$$\alpha = S - (n - ai) \quad (5)$$

25 where i is the number of the critical band, 1 being the number of the lowest frequency critical band, n and a are constants, a is greater than 0, S is the intensity of the convolution processed bark spectrum, and $(n - ai)$ is the allowed function. In this embodiment, n is set to 38 and a is set to 1. This provides satisfactory results with no degradation of sound quality.

30 In this way, the level α is determined, and is transmitted to the divider 26, which applies deconvolution to the level α in the convoluted

region. Accordingly, by carrying out this deconvolution, a masking spectrum is provided from the level α . This masking spectrum becomes the primary allowable noise spectrum. It is to be noted that, while normally deconvolution processing requires a complicated operation, a simple divider
5 26 is used in this embodiment to carry out deconvolution.

Then, the masking spectrum is transmitted to the subtractor 28 through the synthesis circuit 27. Here, the subtractor 28 is supplied with the output of the energy calculating circuit 22 for each critical band, i.e., the previously-described bark spectrum SB, through the delay circuit 29.

10 Accordingly, at the subtractor 28, a subtraction operation between the masking spectrum and the bark spectrum SB is carried out. Thus, as shown in Fig. 12, the portion of the bark spectrum SB having a level lower than the level of the masking spectrum MS is subjected to masking.

BI
cont
15 The output from the subtractor 28 is taken out through the allowable noise corrector 30 and the output terminal 31, and is sent to a ROM, etc. (not shown) in which, e.g., allocated bit number information are stored. The ROM, etc. provides quantizing bit number information for each critical band in response to the output obtained through the allowable noise corrector 30 from the subtractor 28 (i.e., in response to the level difference between
20 energy in each critical band and the output of the allowable noise calculating circuit). The quantizing bit number information is sent to the adaptive bit allocation circuit 18, (Fig. 1) where the spectral coefficients from the DCT circuits 13, 14 and 15 are quantized using numbers of bits allocated to each critical band.

25 The adaptive bit allocation circuit 18 quantizes the spectral coefficients in each band using the number of bits allocated in response to the difference in energy between respective critical bands and the output of the allowable noise calculating circuit. The delay circuit 29 is provided to delay the bark spectrum SB from the energy calculation circuit 22 to take
30 account of delays in the circuits preceding the synthesis circuit 27.

The synthesis circuit 27 synthesizes the minimum audible level curve RC and the masking spectrum MS. The minimum audible level curve is a characteristic of the human sense of hearing, as shown in Fig. 13, and is delivered from the minimum audible level curve generator 32. According to the minimum audible level curve, noise having an absolute level below the minimum audible level curve cannot be heard. Even with the same coding, the minimum audible level curve depends on the volume at the time of reproduction. However, since there is not a great variation in the manner in which a music is represented by, e.g., the 16-bit dynamic range in actual digital systems, if it is assumed that the quantizing noise in the frequency band in which the ear is most sensitive, i.e., in the vicinity of 4 kHz, quantizing noise less than the level of the minimum audible level curve can be regarded as being inaudible in other frequency bands. Therefore, if it is assumed that the system is used such that the quantizing noise near 4 kHz, for a certain quantizing word length, is inaudible, and that the allowable noise level is obtained by synthesizing the minimum audible level curve RC and the masking spectrum MS, then the allowable noise level in each critical band will be the greater of the level of the minimum audible level curve and the masking level. This is shown by the hatched lines in Fig. 13. In the present embodiment, the level of the minimum audible level curve at 4 kHz is matched to the minimum level corresponding to, e.g., quantizing using 20 bits. Fig. 13 also shows the signal spectrum SS.

As explained above with reference to Figs. 3 to 6, in the critical bands where the minimum audible level is selected as the allowable noise level, quantizing bit allocation is performed by dividing the critical band into sub bands. The minimum audible level curve from the minimum audible level curve generator 32 and the masking spectrum MS from the divider 28 are compared with each other in the comparator 35. The result is sent to the synthesis circuit 27. The flag F_{RC} is taken from output terminal 36. For example, in the bands B_{11} and B_{12} of Fig. 13, since the level of the minimum

B1
cont

audible level curve RC is higher than the level of the masking spectrum MS, the minimum audible level curve level RC is selected as the allowable noise level, and the flag F_{RC} is set to 1. Thus, the level of the minimum audible level curve RC for the lowest frequency of the sub bands into which the critical band is divided will be transmitted. As described above, calculation of allowable noise levels for the other sub bands is carried out in the expander.

The allowable noise level corrector 30 corrects the allowable noise level at the output of the subtractor 28 in response to information regarding, e.g., the equal loudness curve from the correction information output circuit 33. The equal loudness curve is a curve characterizing another characteristic of the human sense of hearing. The equal loudness curve corrects sound pressure levels at different frequencies so that they are perceived as sounding as loud as a pure sound at 1 kHz. The equal loudness curve has substantially the same characteristic as the minimum audible level curve RC shown in Fig. 13.

According to the equal loudness curve, a sound in the vicinity of 4 kHz is perceived as being as loud as a sound at 1 kHz having a sound pressure level 8 to 10 dB higher. On the other hand, a sound in the vicinity of 50 Hz must have a sound pressure level some 15 dB higher than a sound at 1 kHz sound to be perceived as sounding as loud. Because of this, the allowable noise level must be corrected using the equal loudness curve to adjust the allowable noise level for the loudness sensitivity of the human sense of hearing.

Additionally, the correction information output circuit 33 may also correct the allowable noise level in response to the difference between the actual number of bits used by the adaptive bit allocation circuit 18 (Fig. 1) to quantize the spectral coefficients, and the target number of bits, which is the total number of bits available for quantizing. The reason why such a correction is made is as follows. There are instances in which there is an

B1
cont

error occurs between the total number of bits allocated by the primary bit allocation process and the target number of bits, which is determined by the bit rate of the compressed digital signal. In such instances, the quantizing bit allocation is made for a second time to reduce the error to zero. For example, if the total number of bits allocated is less than the target value, a number of bits equal to the difference between the actual number of bits and the target number of bits is allocated among the critical bands to provide additional bits. Alternatively, if the actual number of bits is more than the target number of bits, a number of bits corresponding to the difference between the actual number of bits and the target number of bits is removed from the critical bands to remove excess bits.

To correct the actual number of bits, the difference between the actual number of bits and the target number of bits is measured and the output correction information output circuit 33 provides correction data that is used to correct the numbers of bits allocated to the critical bands. Where the error data indicates that insufficient bits have been allocated, an increased number of bits are used per critical band. Conversely, where the error data indicates that excess bits have been allocated, fewer bits can be used in each critical band. The correction information output circuit 33 provides data for the correction value for correcting the allowable noise level at the output from the subtractor 28, e.g., on the basis of information data of the equal loudness curve, in response to the error data. The correction value is transmitted to the allowable noise level correction circuit 30. Thus, the allowable noise level from the subtractor 28 is corrected.

The above-described synthesis processing for the minimum audible level curve may be omitted. In this case, minimum audible level curve generator 32 and synthesis circuit 27 are unnecessary, and the output from the subtractor 24 is subjected to deconvolution at the divider 26, and is transmitted immediately to the subtractor 28.

B1
CONT

Block floating processing and block floating release processing may also be applied in the expander, before and after, respectively, the inverse orthogonal transform (IDCT) processing. In the expander, the logical sum of the absolute values of the spectral coefficients for each block may be taken to determine the block floating coefficient for the block.

In Fig. 14, the input terminal 51 is supplied with quantized spectral coefficients obtained from the output terminal 19 of the compressor shown in Fig. 1. The quantized spectral coefficients is sent to the adaptive bit allocation decoder 52, where the adaptive bit allocation applied by the adaptive bit allocation circuit in the compressor is reversed. The resulting spectral coefficients are sent to the block floating processing circuit 56, where block floating processing is applied to each block of spectral coefficients in each frequency range. Then, the blocks of block floating processed spectral coefficients are subject to inverse orthogonal transform (IDCT, i.e. Inverse Discrete Cosine Transform circuits 53, 54 and 55 in the example of Fig. 14) processing, inverse to the orthogonal transform processing applied by the respective orthogonal transform circuits 13, 14 and 15 of Fig. 1. The outputs from the inverse orthogonal transform circuits 53, 54 and 55 are sent to the block floating release circuit 57, where block floating release processing is applied to each block using the block floating information from the block floating processing circuit 56. The resulting frequency range signals in the respective frequency ranges from the block floating release circuit 57 undergo, by using the synthesis filters 58 and 59, processing opposite to the processing by the band division filters 11 and 12 of Fig. 1, so that the respective frequency range signals are synthesized to provide a single digital output signal. The digital output signal thus synthesized is taken out from the output terminal 60.

It is to be noted that this invention is not limited to the above-described embodiment, but is applicable, e.g., not only to a signal

B1
cont

processing apparatus for an audio signal but also to a signal processing apparatus for a digital speech signal or a digital video signal, etc.

As described above, the apparatus for a compressing a digital input signal according to this invention is adapted to carry out block floating processing of the digital input signal in variable length blocks, and thereafter to implement orthogonal transform processing thereto. In the compressor, by determining the size of each block and the block floating coefficient of the block floating processing in response to the same index, it is possible to reduce a quantity subject to processing, or the number of steps of a program.

Further, in accordance with the apparatus for compressing a digital input signal, when the allowable noise level for each critical band is determined using the minimum audible level, bit allocation is carried out according to the allowable noise level in sub bands obtained by further dividing a critical band, and only a flag indicating this is transmitted. This avoids the necessity of sending an allowable noise level for each sub band. Accordingly, accurate allowable noise levels can be provided without increasing the quantity of auxiliary information transmitted. This way, signal quantity can be improved without degrading the data compression efficiency. In addition, even if the absolute value of the minimum audible level is altered later, compatibility can be maintained.

B1
cancel